



University of Connecticut
OpenCommons@UConn

Honors Scholar Theses

Honors Scholar Program

Spring 5-8-2011

The Politics of Health Outcomes and Income Inequality: A Time Series Cross-Sectional Analysis of County-Level Mortality Rates in the United States

Timothy H. Callaghan

University of Connecticut - Storrs, timothy.h.callaghan@gmail.com

Follow this and additional works at: https://opencommons.uconn.edu/srhonors_theses



Part of the [American Politics Commons](#), [Family, Life Course, and Society Commons](#), and the [Health Policy Commons](#)

Recommended Citation

Callaghan, Timothy H., "The Politics of Health Outcomes and Income Inequality: A Time Series Cross-Sectional Analysis of County-Level Mortality Rates in the United States" (2011). *Honors Scholar Theses*. 183.

https://opencommons.uconn.edu/srhonors_theses/183

Timothy H. Callaghan
Undergraduate Honors Student
University of Connecticut
53 Carman Road
Manchester, CT 06042
timothy.callaghan@uconn.edu

“The Politics of Health Outcomes and Income Inequality: A Time Series Cross-Sectional Analysis of County-Level Mortality Rates in the United States”

Abstract:

Health inequalities are pervasive in the United States today. Despite social epidemiologists frequently citing political and economic factors for this variance, political science has largely ignored these issues. Given this gap in the literature and the importance of the issue, more research is clearly needed to better understand the political and economic causes and implications of these health disparities. This study analyzes the topic in depth, examining how income inequality, which is believed to be a key factor in explaining health inequalities, is related to mortality rates at the county level. Examining aggregate data from all US counties from 2000-2006, this study provides the first ever county-level time series, cross-sectional analysis of the topic. Furthermore, I advance the efforts of social epidemiologists by placing health inequalities in a political theoretical framework. Ultimately, this paper shows that income inequality is positively correlated with mortality in US counties, and that several additional political economic variables, such as county level Republican deviation in presidential elections, level of unemployment, and aggregate personal income, also have a significant effect on health inequalities.

Introduction

Health inequalities are pervasive in the United States today. The political and economic causes of these inequalities however are poorly understood. With social epidemiological studies showing minorities and those with lower socio-economic standing up to “five times more likely to be hospitalized for diabetes, twice as likely to be hospitalized for and to die of heart disease, three times more likely to die of stroke and twice as likely to die of cancer” than their more well off counterparts, extensive research is needed to understand the root causes of these differences in health outcomes (Beyers and Brown 2008, 1). This thesis examines these key factors in depth, ultimately determining that increased income inequality, which researchers believe is the key factor leading to differences in health, leads to increased mortality at the county level and further that several additional political economic variables, including level of unemployment and aggregate personal income, also impact health inequalities.

The role of income inequality in political outcomes has begun to receive considerable attention from political scientists (Bartels 2008; Gelman 2008; Kelly 2009; McCarthy et al. 2006; Page and Jacobs 2009; Soss and Jacobs 2009). Furthermore, it has been identified as having an important role in international (Theodossiou and Zangelidis 2009), national (Wilkinson and Pickett 2008) and local (Franzini and Spears 2003, Brodish et al. 2000, and Auger et al. 2009) health outcomes. Through these studies, it is becoming increasingly clear that income inequality has pervasive effects; however income inequality and health outcomes remain under-theorized and under-analyzed.

Income Inequality and Health Outcomes

Health inequalities, defined as the “term used to designate differences, variations, and disparities in health achievements of individuals and groups,” are prevalent in the United States today and detailed research is needed to understand their causes and consequences (Kawachi et al. 2002, 647). These inequalities are most pronounced when examining mortality. For example, Beyers and Brown (2008) recently found that disadvantaged Americans die up to 15 years prior to their more well-off counterparts. These discrepancies in health have been linked to income inequality in numerous studies at all levels of analysis and have been explained using three competing theories; psychosocial determinants of health, neo-material determinants of health, and individual income. These theories have led to disagreements in the field; with researchers debating why income inequality negatively affects health outcomes (Lynch et al. 2000, 1200-1202).

The first of these theories, the psychosocial interpretation of income inequality, argues that an individual’s personal sense of where they stand on the social ladder relative to others can actually affect health, with those who feel inferior exhibiting stress which leads to poorer health outcomes (Lynch et al. 2000, 1201). “Simultaneously, perceptions of relative positions and the negative emotions they foster are translated outside the individual into antisocial behavior, reduced civic participation and less social capital and cohesion within the community” (Lynch et al. 2000, 1201). As poor health in this theory is the result of how individuals view themselves as compared to those around them, areas with less inequality lead to better health outcomes, with individuals less

likely to experience the “anxiety, insecurity, and depression” seen when inequalities exist (Marmot and Wilkinson 2001, 1233).

While the psychosocial theory has been studied extensively and many researchers in the field agree with it, others have disagreed, arguing that inequality’s impact on health outcomes is better explained through the neo-material theory. Proponents of the neo-material theory argue that health inequalities are caused by the different experiences individuals’ face, with individuals who lack material resources experiencing poorer health (Lynch et al. 2000, 1202). These differences in resources, as seen through differential access to “education, health services, transportation,...availability of food, quality of housing, [and] occupational health regulations” are the result of income inequality and lead to differences in health (Lynch et al. 2000, 1202). Unlike psychosocial theorists who argue that inequality leads to poor health outcomes based on the negative feelings inequality creates, neo-material theorists instead argue that these poor health outcomes are caused by the tangible disadvantages faced by those with increased inequality. Proponents of the theory don’t deny that there may be negative emotions that inequalities create, however they argue that the lack of material wealth as seen through decreased access to resources, leads to the poor health, not the emotions of those impacted by the inequality (Lynch et al 2000, 1202-1203).

Finally, although the neo-material and psychosocial theories are the most commonly cited as the reason for income inequalities’ impact on health, a third theory known as the individual income theory also exists. In this theory, proponents argue that it is not a lack of resources or perceived social position that leads to poor health, but instead that increased inequality leads to more poor individuals, and in turn, poorer health.

Noting that “a society with greater income inequality will have a higher percentage of people with low incomes, and that this higher prevalence of poor people accounts for the relation with poor health”, this theory is the most controversial of the three (Marmot and Wilkinson 2001, 1233). By asserting that increased poverty alone is responsible for differences in health, some researchers have argued that this theory ignores many of factors leading to poor health seen in the psychosocial and neo-material theories (Lynch et al. 2000 1201; Marmot and Wilkinson 2001, 1233). That said, this theory has proven itself effective in explaining individual level analyses and for that reason is important to include in this analysis as well (Lynch et al. 2000, 1200). Ultimately, although there continues to be disagreement about why inequality negatively affects health outcomes (Lynch et al. 2000, 1200-1202), there is little doubt that a relationship exists between the variables and that further investigation of the topic is needed.

With over 150 peer-reviewed analyses into the relationship between income inequality and health outcomes over the past thirty years, the relationship has been examined in several contexts and numerous methods for measuring income inequality have been used (Wilkinson and Pickett 2006, 1769). These differences in methodology have had a drastic impact, with variable choice accounting for most of the differences in results seen in the field. While international comparisons of income inequality must use data available across countries, often leading to more general studies, lower level analyses like the studies of Franzini and Brodish at the county level allow for more detailed analyses using data from the US Census (Brodish et al. 2000, 387 and Franzini et al. 2001, 376). These differences are then exacerbated by the choice of health variables within studies, with some researchers such as Wilkinson choosing to focus on mortality,

while others like Theodossiou and Zangelidis have focused on mobility and self health assessments (Theodossiou and Zangelidis 2009, 230 and Wilkinson and Pickett 2008, 699). Together, the various combinations of income and health variable chosen in studies of health inequality have led to different results by researchers, making it difficult to conclusively determine the effects of income inequality on health.

These difficulties in determining how the variables are related have been made even more challenging due to the differing methodologies chosen by researchers in examining inequality, as there is no set method for measuring the variable (Franzini et al. 2001, 376). While some studies have analyzed how income inequality relates to health using the Gini coefficient, others have chosen to use the 90/10 ratio or even econometric equations. While each researcher has justified their choice of inequality measure with Brodish for example arguing for the use of the Gini coefficient as it is a common method for measuring inequality, and Theodossiou and Zangelidis arguing for the use of econometric equations as extraneous variables can be controlled for, these differing methodologies have made it more difficult to compare the results obtained between studies (Brodish et al. 2000, 387 and Theodossiou and Zangelidis 2009, 231). That said, with 131 of the 168 analyses into the field of health inequalities showing that income inequality is at least partially responsible for differential health outcomes, it becomes clear that income inequality is strongly correlated to health outcomes, regardless of the methodology chosen (Wilkinson and Pickett 2006, 1769).

This idea becomes clear through an examination of the prominent literature in the field, with almost every major article showing at least some correlation between increased inequality and poorer health outcomes. Although exhibiting differential

outcomes based on the level of analysis, with much stronger correlations seen within countries as opposed to between them, studies consistently show a positive correlation between the variables (Wilkinson 1997a, 591). Furthermore, detailed studies using multi-level analyses like Wilkinson and Pickett's 2008 study of income inequality and mortality at both the state and county level, have shown that relationships are consistent between levels of analysis, finding that "counties at the same level of income will have lower mortality if they are from more...equal states" (Wilkinson and Pickett 2008, 701).

Other studies have gone even further, examining not only whether income inequality is related to health at various levels of analysis, but also examining the impact of race and social position on the variables. These studies have helped to further researchers understanding of income inequality and health outcomes, for example showing that minorities were likely to die 3-4 years earlier than members of the dominant race (Franzini and Spears 2003, 1855). Attributed to the stress of being a minority and the stigmas and stereotypes associated with that designation, these studies have shown a direct relationship between inequality and poorer health outcomes (Franzini and Spears 2003, 1855). Although important, other prominent studies have shown that the impact of inequality on health can be somewhat subjective, with "individuals who feel that they are doing better compared to their near equals enjoying better health" (Theodossiou and Zangelidis 2009, 234). This outcome has led some to believe that inequality is caused not by some external factor, but instead by how individuals perceive themselves, with health determined by whether someone feels they are superior or inferior to those around them (Theodossiou and Zangelidis 2009, 234). While different in their interpretations, these studies have consistently shown a strong relationship between income inequality and

poorer health outcomes and have served as the basis for my background knowledge in my analysis of the topic.

While the vast majority of studies into the field of health inequalities have shown a positive correlation, it is important to acknowledge the fact that some studies have unexpectedly produced negative results. One such study, analyzing income inequality and health outcomes in the communities of Quebec, found that all cause mortality for the elderly, alcohol related mortality and overall mortality were lower in areas with higher income inequality (Auger et al. 2009, 441). Similar to results obtained in other studies reporting a negative correlation, the authors explain their result noting “that high income inequality may be a characteristic of generally affluent neighborhoods with less variation in household income” (Auger et al. 2009, 442). Although this explanation would account for the unexpected results of this study, based on countless studies exhibiting positive results it is far more likely that this result was caused by either poor variable choice, or outside factors like the presence of socialized medicine (Wilkinson and Pickett 2006, 1778). This reasoning is promoted by Wilkinson who argues that poor variable choice, seen by either using areas of analysis too small for the effects of inequality to appear, or through variables which fail to act as good measures of inequality, are the likely cause of many of the negative correlations seen between income inequality and health outcomes (Wilkinson and Pickett 2006, 1778). As over seventy five percent of studies into the field of income inequality and health outcomes have shown a positive correlation between the variables, and most that have not attributable to poor variable choice, the vast majority of evidence shows that income inequality does have an impact on health outcomes.

After an in depth examination of the literature, it becomes clear that income inequality has been consistently linked to health outcomes, with increased inequality leading to poorer health internationally, nationally and at the local level. With over seventy five percent of the literature in the field exhibiting a positive correlation and the studies that do not easily attributable to poor variable choice, this relationship has proven consistent over the past thirty years and can be trusted. Historically explained using the psychosocial, neo-material, or individual income interpretation, these theories will provide crucial background knowledge as I attempt to determine the exact nature of the relationship between the variables. Ultimately, this literature review has shown that with proper variable choice, as well as a complete understanding of the relationship between the various health inequality interpretations, my thesis will be able to give a complete view of health inequalities; drawing on the work of others and coming to new conclusions based on my results. Using an extended period of analysis, a larger sample size than previous studies, and including political variables previously ignored, this study will provide a significant addition to the field.

Data and Variables

This analysis will study how income inequality affects mortality at the county level, examining the relationship in all 3,141 US counties from 2000-2006. Although requiring a large amount of data collection, a comprehensive analysis of the topic at the local level is critical in gaining a better understanding of the issue. With 83% of studies at the international level, 73% of studies at the national level, and only 45% of studies at the local level finding a positive correlation between income inequality and health outcomes,

a detailed analysis at the county level will help to determine whether or not a correlation between the variables actually exists at the local level (Wilkinson and Pickett 2006, 1774). Further, by examining the topic from 2000-2006 as opposed to a single year as seen in previous studies such as Wilkinson and Pickett (2008), Brodish et al. (2000), and Franzini et al. (2001), this study will be able to show how changes in income inequality have affected health over time; providing much greater detail than previous studies which could only determine if a correlation existed. Creating a seven year period of analysis, this time frame was chosen based on the availability of both the mortality and income inequality data, as these were the only years that the two datasets overlapped. Ultimately, by examining this topic in all US counties and over an extended period of time, this analysis will provide a far clearer picture into health inequalities than previous studies and will help to determine if income inequality affects mortality at the local level.

The dependent variable in this analysis, healthcare outcomes as seen through mortality, was obtained using all cause mortality data made available through the compressed mortality files in the Center for Disease Control's Wonder database. Income inequality on the other hand, the key independent variable in the study, was obtained using income data available through the US Census. Although not the focal point of the study, other variables are used as well, serving as controls in my attempt to understand the impact of income inequality on health. These other variables, including total county population, unemployment, number of individuals on Medicare and number of individuals below the poverty line in each county, serve as controls in the study, helping to remove the impact of other variables on health and allowing for a better understanding of the relationship between income inequality and mortality.

In discussing how income inequality affects health in depth, it is important to discuss the variables chosen for the study and how they are measured. In this analysis, the dependent variable is health outcomes measured using mortality from all causes of death. While some studies have used other methods for determining the affects of inequality on health, most notably mobility, self-health assessments and even psychological well-being, by using mortality as a measure of health outcomes, this study removes the bias associated with attempting to determine the relative well-being of an individual (Theodossiou and Zangelidis 2009, 231). Obtained from the CDC's Wonder database which contains mortality data broken down by year, county, and type of death; using all-cause mortality in this analysis has allowed me to provide a picture of the entire relationship between the variables (Centers for Disease Control and Prevention 2009). Recorded based on an individual's place of residence at the time of their death as opposed to where they actually died, and measured as the number of deaths per 100,000 individuals living in a given county, this variable can reveal stark differences in health outcomes (CDC Wonder Customer Support, October 6th, 2010). With Chattahoochee County in Georgia for example showing a mortality rate of 215.1 deaths per 100,000 individuals, and Iron County Michigan showing a rate of 1837.0 for the same year, an 854% difference in mortality, it becomes clear that mortality acts as a good measure to study differences in health outcomes between counties (Centers for Disease Control and Prevention 2009).

While this analysis has shown that mortality acts as a good measure of health outcomes, it is important to acknowledge that the mortality dataset from the CDC is imperfect and is missing data points from some counties, a fact which must be controlled

for in my analysis of the topic. With data points from several counties coded by the CDC as suppressed, these errors have needed to be accounted for in the dataset in order to ensure that my study provides a complete analysis of the variables at the county level. Coded as suppressed if there are less than 6 deaths in a county in a given year, the CDC has chosen not to report counties with low death counts out of concern that a given death could be traced to an individual (Compressed Mortality File 1979-1998 and 2000-2006, 2009). While understandable as the CDC should first and foremost protect the privacy of the individuals used to create the dataset, these suppressed data points have been coded as missing data in the study; a fact which has decreased the accuracy of the analysis. That said, as only a select few counties have suppressed data, this study still uses the vast majority of the 3,141 US counties in the analysis, a sample size large enough that I feel confident that these missing data points will not have a significant effect on my results.

Outside of mortality, the other main variable used in this analysis of the topic was income inequality, the key independent variable in the study. With no common measure for assessing the variable, researchers have tried numerous methods in their attempts to operationalize the variable (Franzini et al. 2001, 376). With researchers using methods as wide ranging as the Gini coefficient (Brodish et al. 2000, 387), geocoding (Franzini and Spears 2003, 1851), the Robin Hood index (Franzini et al. 2001, 376), and even econometric equations (Theodossiou and Zangelidis 231), there was no standard to use in this analysis of the topic, and for that reason I will use a new method to measure income inequality. Created by Dr. Jeffrey W. Ladewig using county level data from the US Census, this new method of calculating income inequality is measured as “the ratio of...mean income to median income” (Ladewig 2010, 39). Using both aggregate personal

income and median household income data available through the US Census, this somewhat simplistic method has proven effective in measurements of income inequality as it relates to party polarization and should prove effective in this analysis as well (Ladewig 2010, 39). As mean income rises relative to median income, a society becomes more and more unequal as a larger portion of income is held by a small group (Ladewig 2010, 6). Using this method, the closer the ratio of mean to median income is to one; the more equal a society is, with values above one showing at least some income inequality. Although not necessarily a perfect measure of income inequality, this ratio should prove effective in this analysis of the topic and I expect to find that the more income inequality there is in a given county, the higher the mortality rate will be.

Other variables were used in the analysis as well to act as control variables in this study of income inequality and health outcomes. By performing the analysis at the county level, the variables available to the analysis were limited, restricting the study to variables for which comprehensive county level data is available. Despite this limitation, this study uses several control variables which could themselves be responsible for differences in health outcomes, ensuring that the analysis provides an accurate picture into the relationship between income inequality and health outcomes.

The first of these variables, population, is measured as the natural log of the total population in each county. With vast differences in population among counties in urban and those in more rural areas, accounting for these differences is critical in gaining an understanding of how income inequality affects mortality. By using the natural log of population as opposed to just the crude measure of population, the study is able to control for any diminishing returns that may be seen as the size of the population increases. For

example, it is critical to control for the fact that differences in mortality rates may be far greater between populations of 100 and 100,100 individuals than they would be between 1,000,000 and 1,100,000 individuals. As densely populated areas tend to have more doctors and hospitals, and therefore shorter commutes to receive needed care, I expect to find that the higher the total population of the county is, the lower the mortality rate in the county will be.

Another important control variable will be the number of African Americans living in each county in the United States. As previous studies have found that being black is attributable to 3 to 4 years of lost life, it is important to include this variable in the analysis to serve as both a comparison with previous studies, as well as to ensure that a complete understanding of health inequalities is obtained (Franzini and Spears 2003, 1855). Studied previously by Franzini and Spears (2003) as well as Byers and Brown (2008), this study will act as a comparison to the results obtained in their studies and I expect to find that the more African Americans live in a county, the higher the mortality rate will be.

The third control variable to be used in the study will be the number of females living in each county. Serving as a measure for gender, by using the number of females per county this study will be able to see what impact gender has on mortality within a given county. Using differences in life expectancy as a guide, which consistently show females living longer than males, as well as studies showing that females are more conscious of their personal health, I expect to find with this variable that the more females there are in a county, the lower the death rate will be, attributable to the county having fewer males who are likely to die early.

The next control variable to be used in the analysis will be the number of individuals in the county on Medicare. Collected as the number of individuals on Medicare A (which covers hospital costs) and Medicare B (which covers traditional medical insurance for doctors visits), the number of individuals on Medicare is a critical variable to collect in any analysis of mortality (“Getting Started” 2010). As individuals on Medicare are elderly or the very sick (e.g., in the case of those receiving dialysis treatment), the number of individuals on Medicare serves as a variable for those with the highest probability of death in each county. Following the simple logic that those with a higher probability of death will die more frequently than those with a lower probability of death, I expect to find that counties with a large number of individuals receiving Medicare will have a higher mortality rate than those with fewer Medicare patients.

The fifth control variable to be used in the analysis will be the income level in each county, measured using personal income for the county as a whole. Although difficult to choose a specific measure with so many income variables available, this analysis will use mean income per housing unit at the county level, a common income measure which will allow for comparisons to previous studies. While the focus of this study is income inequality, it is critical to have an understanding of how income impacts mortality as well, with income playing a critical role in the formulation of the income inequality measure. As increased income has been linked to increased access to material resources, as noted in the neo-material theory, I expect to find that as the aggregate income within a county increases, the mortality in the county will decrease.

Another important control variable will be the number of unemployed individuals living in each county. As unemployment is linked to inequality, with employed and

unemployed individuals experiencing substantial differences in access to resources, it is important to include this variable in studies of income inequality in order to ensure that an accurate depiction of income inequality is seen. With unemployment inextricably linked to income inequality, and this paper's expectation that as income inequality increases, so too will mortality, I expect to find that as unemployment increases, mortality will increase as well.

Next, the level of poverty within each county will be used as a variable to control for variations in mortality caused by the presence of poverty as opposed to differences in income inequality. The impact of poverty is a prominent feature of both the individual income and neo-material theories on income inequality, and for that reason I felt that the variable was a necessary addition to my analysis. Measured as the percent of the population in each county below the poverty line, with this variable I expect to find that as the percentage of impoverished individuals within a county increases, the mortality rate will increase as well.

Another new variable included in this analysis of income inequality, was the number of older individuals living in each county. While my analysis also includes a control variable for the number of individuals on Medicare A and B living in each county, the Medicare variable fails to account for older individuals who may not qualify for both Medicare types A and B or who are not on Medicare at all. While all individuals automatically qualify for Medicare A (hospital insurance) so long as they have "paid Medicare taxes for at least ten years", Medicare B (medical insurance) is only available for those who are willing to pay for it through a monthly reduction from their social security ("Support Center" 2001). For that reason I chose to include this new variable

which measures the number of individuals within each county over the age of 65. As individuals over the age of 65 are more likely to have chronic diseases and have poorer health than the rest of the population, with this variable I expect to find that the mortality rate will be higher in counties with larger populations of citizens over the age of 65.

Similarly, this analysis also controls for the mean age group within the population as a whole. While including the number of individuals over the age of 65 in the analysis is important, a far more telling variable should be the mean age of the population in each county. Broken down into 18 age groups by the US Census, this variable is a weighted average of the 18 age groups to determine the average age of the citizens in each county (Ladewig and Callaghan 2011, 9). Following the same logic as the variable controlling for the number of individuals in each county over the age of 65, I expect to find that the older a given population is on average, the higher the mortality rate will be.

Next, two variables will be used in the analysis “that account for the geographic attributes” of each county (Ladewig and Callaghan 2011, 11). First, the actual size of the county will be measured using a control variable that accounts for the land area of every county using units of 100 miles. With larger counties likely having more rural areas and therefore longer commutes to medical care, I expect to find that as the size of the county increases, the mortality rate will increase as well. Besides county size, this analysis will also control for another important geographic feature, the population density within each county. Measured as the product of total population and land area, I expect to find that mortality will be lower in densely populated areas, with more doctors and hospitals likely in areas with a high density of citizens (Ladewig and Callaghan 2011, 11).

Finally, in order to control for social cohesion, a key control variable in any analysis of income inequality, this study will measure “the county-level [deviation of the] percent of two party support for the Republican candidate for President from the national two-party average of the Republican candidate for President” (Ladewig and Callaghan 2011, 10). Although an imperfect measure of social cohesion with many factors affecting voter choice, this variable is commonly used to represent the ideology of a given area and should be an effective control variable in my analysis. With Republicans generally more wealthy than their liberal counterparts and thus potentially better able to get the care that they need based on the neo-material and psychosocial theories, I expect to find that the more Republican a county is, the lower the death rate will be. That said, with Republicans generally in opposition of government intervention into health, this relationship is extremely complex and more research will be needed in the future to better understand this variable.

Ultimately, while it is quite difficult to accurately operationalize and measure the variables used in any study, by using mortality to measure health outcomes, the ratio of mean to median income to measure income inequality, and several control variables to ensure that an accurate measure of the effect of income inequality on health is seen, this study provides an analysis into how the variables are related. Using regression models to see if a correlation exists, and then comparing changes in the variables over time, this analysis shows how income inequality affects health.

Results

To test these theories and determine the exact relationship between income inequality and mortality at the county level, a least squared regression for time-series cross-sectional data was run on all variables included in the analysis. The model allows for clustering by county. In this way, “the model provides coefficient estimates that are robust to cross-sectional heteroskedasticity and autocorrelation” (Ladewig 2010). The results of this model can be seen below in table 1 and all results were determined to be highly significant and robust.

[Table 1 Here]

In analyzing the results of the analysis, the regression model found the relationship between income inequality and mortality to be statistically significant and in the expected direction, supporting the assertion that as income inequality in a county increases, so too does mortality. While this result differs from the slight majority of previous studies into income inequality at the local level, as 55 percent of local level analyses have found a negative correlation, this analysis was far more comprehensive (Wilkinson and Pickett 2006, 1774). Using a seven year period of analysis in all US counties, as opposed to the small area, single year analyses seen in many of the local-level studies producing negative results, this study provides a far more complete view of the relationship at the county level and can be trusted.

Furthermore, by using mean over median income to measure income inequality, this study provides a new method for measurement of the variable in future analysis. As so many other methods for measuring inequality have been inconsistent and are complicated to work with, using this new measure, which has been effective in studies

related to party polarization, could help to advance understanding in the field. Ultimately, by using an extended period of analysis, a large sample size, and a new and effective measure for income inequality, this study lends support to the idea that income inequality is positively correlated with mortality at the county level in the United States.

Beyond examining income inequality, this study also included several control variables in order to provide a complete picture of the relationship between income inequality and mortality. The first of these variables, the natural log of county population, showed a strong negative correlation in the analysis. Supporting the assertion that as the size of the population increases mortality decreases, this result matched the outcome that was expected for the variable. This result is likely the product of better access to care in more populated areas, with individuals in less populous, rural areas forced to go farther for medical care.

The second control variable used in the analysis, the percent of African Americans living in a county, was significant and positively correlated with mortality. Showing that the more African Americans there are in a county, the higher the mortality rate, this result matched the expected outcome of this paper and the previous work of Franzini and Spears (2003). Attributable to the disadvantages faced by minorities such as lower incomes, limited health coverage and any potential feelings of inferiority, this outcome helps to show how inequality affects health.

The next variable used in the analysis, the number of females in a county, served as a measure of gender and was found to be significant and showed a strong negative correlation. Confirming my hypothesis that more females would lead to fewer deaths as

females have longer life expectancies, the model found that the more females there were in a county, the lower the mortality rate was.

The fourth control used in the analysis of income inequality and health outcomes, number of individuals on Medicare, serves as a measure for the elderly and those that are most likely to die. Finding a strong positive correlation in the regression analysis, this variable showed that the more people in a given county were on Medicare, the higher the mortality rate. Matching the expected result for this analysis, this variable shows that while income inequality is the main factor impacting mortality, age and health of the individuals in the county does play an important role. Examining the results of the analysis in depth, the number of individuals in the county on Medicare proved to be one of the best control variables used in the analysis, with a large coefficient showing that the more individuals in the county there were on Medicare, the higher the death rate in the county would be.

The next control variable used in the analysis, poverty, which was measured as the percentage of the population in each county under the poverty line, was found to have a strong positive correlation with mortality. Matching my expected result for the variable, this analysis found that as the percentage of individuals in counties under the poverty line increased, the mortality rate increased as well. One of the critical components of both the neo-material theory which argues that decreased access to resources as a result of poverty leads to higher mortality, as well as the individual income theory which argues that increased inequality leads to increased poverty and, in turn, poor health, this result supports the premise of both theories and shows the importance of poverty on health outcomes.

The sixth control variable used in the analysis, unemployment, was found to be significant and have a positive correlation with mortality. Showing that as the number of unemployed individuals within a county increases, so too does mortality, this result proved my hypothesis. Although the opposite of the result obtained in my preliminary analysis of the topic last semester in which I found a strong negative correlation, this final analysis includes several new control variables such as poverty and the number of individuals in the county over the age of 65. These new variables account for a large portion of the unemployed individuals in the United States, making the unemployment variable far more precise and ensuring that an accurate picture into the relationship between income inequality and health outcomes is seen.

The next control variable examined in the analysis of income inequality and health outcomes was mean age group in each county, and was found to be highly significant and negatively correlated. Serving as a measure for the average age of the population in each county, this result did not match what I expected to see, showing that as mean age group for counties increased, mortality in the counties decreased. While this result was unexpected as it would seem counter-intuitive that fewer deaths would occur in counties with older populations, this outcome could be the result of other control variables (number of individuals over 65 and number of individuals on Medicare) “controlling for the rest of the variance in mortality rates” (Ladewig and Callaghan 2011, 13). Furthermore, as older age groups tend to be higher up in the social hierarchy and have higher incomes, this result lends support to both the neo-material and psychosocial theories which argue that income and social position are critical in determining health outcomes.

While mean age group was an important addition to my analysis as age is a critical factor in mortality, a far more important control variable labeled 'Population Old' in table 1, which represents the number of individuals in each county over the age of 65, was included in this analysis. Although including the mean age of the population is important, individuals over the age of 65 should be the most likely to die in any given county and through my analysis, I found that to be the case. With my analysis showing a strong positive correlation for individuals over 65, this analysis found that as the number of old individuals in counties increased the mortality rate in counties increased as well.

Next, this study included two variables to control for the geographic attributes of the county which could have had a major impact on mortality rates in counties around the country. With the first variable, land area in units of 100 miles, I found a strong negative correlation, showing that as the size of the county increased, mortality rates decreased. While I initially hypothesized that increased county size would lead to increased mortality as individuals would have longer commutes to medical care, the hypothesis is incorrect and further research is needed to understand why larger counties experience lower mortality rates. With the second variable used to control for geographic attributes, population density, I found a strong negative correlation as well. Indicating that more densely populated areas have lower mortality rates, this result matched my hypothesis for the variable. As densely populated areas should have more hospitals and doctors, individuals in these areas are more likely to get the care that they need in a timely fashion, thus decreasing mortality rates.

Finally, county level deviation in percent vote for republican candidates in presidential elections from the national average was examined, serving as a measure of

both social cohesion and belief in government intervention into health. While I admit that this is an imperfect measure, with many factors other than views on healthcare going into an individual's vote decision, it serves as an effective measure of a counties' belief in government programs like Medicare and Medicaid which are supported by Democrats and opposed by Republicans. In analyzing the result of this variable, percent vote for Republicans was found to be significant, and negatively correlated with mortality; showing that the more Republican a county voted as compared to the national average, the healthier it was. While this result was proven statistically significant, it may at first seem unexpected as government programs, which work to decrease income inequality and, in turn, mortality are often opposed by Republicans. That said, as Republicans receive the government program support no matter what party they vote for, Republicans benefit from the programs even if they do not agree with their implementation. Furthermore, as Republicans typically have higher incomes than Democrats, perhaps they have better access to healthcare following the logic of either the neo-material or psychosocial theories.

An Alternative Model

While using table 1 for my analysis of the topic proved quite effective, showing a strong positive relationship between income inequality and health outcomes, as well as revealing strong correlations for all control variables, the regression analysis fails to include one crucial variable, income.

[Figure 5 here]

Income is a critical component in the creation of the income inequality measure, created using the difference between the mean and median income in counties and I felt that it needed to be included in my analysis of the topic. This inclusion is supported by figure 5, which shows a scatter-plot of income and income inequality revealing a strong positive correlation between the variables. By including income in my analysis of the topic, I will be able to ensure that differences in mortality between counties are caused by income inequality as opposed to income itself. For that reason, this analysis of income inequality and health outcomes also includes table 2, containing all of the same variables as table 1, except also including income measured as the mean income per housing unit in all US counties.

[Table 2 Here]

After re-running the regression using all of the variables in table 1 as well as income, the new coefficients revealed many interesting characteristics about the study. First, as income inequality continued to show a strong positive correlation, this analysis is able to show that the differences in mortality were caused by income inequality as opposed to just differences in income. With that in mind however, it is important to acknowledge the impact of income on mortality, with mean housing unit income revealing a strong negative correlation, indicating that the higher incomes within a county are on average, the lower the mortality rate is.

Outside of the critical variables income and income inequality, the inclusion of income in the regression also had an impact on several of the control variables used in the analysis. While some variables showed similar correlations to those seen in the regression run without income, with number of African Americans, number of individuals on

Medicare A and B, natural log of the population, county Republican deviation from the national average, the number of individuals over the age of 65, and land area in units of 100 miles all showing strong correlations in the same direction as in the regression run without income, the inclusion of income had unexpected effects on other control variables.

By including income in the regression to determine the impact of income inequality on mortality, 4 key control variables lost significance of any kind. The number of females living in a county, the number of individuals below the poverty line, the number of unemployed individuals and the mean age group all lost significance. Furthermore, population density, which is simply the product of population and land area, also lost significance although to a lesser extent than the other variables, going through a drop from $p < 0.01$ to $p < 0.05$, making the result somewhat significant. While this result may be surprising at first, with 5 control variables losing at least some significance with the addition of 1 new control variable, this outcome can be explained by the fact that the income variable accounts for the correlation seen between mortality and the other variables. With this outcome, although it may be intuitive to think that the non-significant control variables are not important, in reality it merely shows that the inclusion of income masks the effects of those other variables on mortality. In the end, by including both table 1 without income and table 2 with income, this analysis is able to give detailed insight into how income inequality impacts health outcomes and how numerous control variables also play a crucial role.

Validity Test

While this analysis of income inequality and health outcomes uses numerous control variables to ensure that the study provides accurate insights into how income inequality impacts mortality, one critique often made about studies of these variables is that the results seen are caused by extreme outliers as opposed to variations in counties across the country. For that reason, I have also included several histograms of my results in order to show that extreme differences in variables, for example the differences in mortality rates between Iron County Michigan (1837.0/100,000 individuals) and Chattahoochee County in Georgia (215.1/100,000), represent a wide variation between counties as opposed to just a few outliers.

[Figures 1 and 2 Here]

This wide variation can be seen clearly through analyzing the two main variables used in my study, mortality which can be seen in figure 1 and income inequality which can be seen in figure 2. Both histograms show normal distributions with the mortality variable centered on 1000 deaths/100,000 individuals and income inequality centered slightly above 1.5. With both variables showing normal distributions and clearly illustrating a wide variation in results as opposed to extreme outliers, these histograms are able to overcome potential challenges to the analysis' legitimacy.

[Figures 3 and 4 here]

Although the two main variables used in the analysis showed normal distributions with a wide variance helping to legitimize the analysis, I also chose to create histograms for two key control variables, mean housing unit income and the percentage of the county

under the poverty line which showed unexpected results. Both variables illustrated positive/right skewness, with the mean income in figure 3 going as high as over 150,000 dollars and the percentage of individuals below the poverty line going as high as 40%. While this skewness was unexpected, showing that some of my control variables had outliers, the presence of outliers is less of a concern in controls and should have only a minor impact on the study as a whole. Ultimately, with the key variables in my analysis revealing normal distributions and only a few of my control variables showing unexpected outliers, this paper is able to overcome a major criticism of previous works and reveals important insights into how income inequality impacts mortality.

Discussion and Conclusion

This analysis of income inequality and mortality in all US counties from 2000-2006 has shown that income inequality and mortality are positively correlated, with increasing inequality leading to higher mortality rates. Using several common control variables including county population, number of African Americans, and number of individuals below the poverty line, as well as a less common variable, Republican deviation from the national average, this study found all variables to be significant and strongly correlated with mortality. These control variables increased the validity of the study and helped to provide an accurate picture into the relationship between income inequality and health outcomes. Furthermore, with the addition of income to the analysis, income inequality continued to have a strong positive correlation with mortality, although some control variables lost significance. Ultimately, this analysis shows that income

inequality is a critical determinant in health outcomes in US counties and that even when controlling for other variables, the outcome remained significant and robust.

Although this study cannot prove causation, it does provide some insight into the three theories on how income inequality impacts health. First, as this analysis found a negative correlation between income and mortality, this study would suggest that the individual income theory, which has been effective in studies at the individual level, is also accurate at the aggregate level. With the individual income theory arguing that differences in income and more specifically the presence of poverty leads to poor health outcomes, the positive relationship between poverty and health outcomes also supports this theory (Marmot and Wilkinson 2001, 1233). Next, with this study showing a positive correlation between unemployment and mortality, this study would suggest that the neo-material theory is also at least somewhat accurate. Arguing that decreased access to resources leads to poorer health, if this theory were accurate I would expect to see the unemployed, with extremely limited access to resources, experiencing poorer health than the rest of the population, which in this analysis I did. Finally, with this study showing that an increased number of African American living in a county is associated with higher mortality, this study lends support to the psychosocial theory. With proponents of the psychosocial theory arguing that minorities and blacks in particular should experience poorer health due to “the...stress of being black in America, [including] the relative economic deprivation of blacks and the discrimination encountered in every day life”, and this study finding that African Americans do in fact have increased mortality, this analysis supports the psychosocial interpretation (Franzini and Spears 2003, 1855). While admittedly some of these results lose significance with the addition of the income control

variable to the analysis, these inferences are accurate based on the results in table 1 and lend at least some support to the arguments made in all three theories. As all three theories are at least partially supported by the data, perhaps a new theory is needed that encompasses the key elements of all three. That said, a true determination of which theories are correct is outside the scope of this paper, and the examination of the theories above can do little more than suggest the aspects of each theory that this paper's findings support.

While this study provides many advantages over previous studies, it is important to acknowledge that there are still shortcomings in the analysis that will need to be addressed in future research of the topic. Most importantly, additional control variables could be added to the model. While the final version of this study includes several new variables which were not included in my initial analysis including mean age group, poverty, and geographic attributes, there are many other control variables that could add to the strength of the paper. One such variable that would be useful in a more in-depth analysis of the topic would be the number of individuals on Medicaid in each county. Including a Medicaid variable would help to distinguish between the poor with medical coverage and the poor without it, a critical distinction that currently is not controlled for in my paper. Another useful control variable to add would be the average education level of the citizens living in each county. With personal health inextricably linked to personal knowledge of health and wellbeing, including a variable on the education level within counties could be useful in future analyses of the topic. It would allow researchers to distinguish whether increased mortality is attributable for example to poverty itself, or if counties with lower levels of education (which could result in increased poverty) actually

cause poorer health outcomes. Finally, adding an interactive term between income and income inequality may provide additional analytical leverage, because income and income inequality together affect mortality.

Outside of variables to be added in future analysis of the topic, another issue that will need to be addressed in future studies of the topic will be aggregate personal income, seen in this analysis as mean housing unit income. With tables 1 and 2 showing the drastic impact of income on the analysis, with 5 key control variables losing at least some significance with the addition of income to the regression, further analysis of income's impact on not only mortality and income inequality but the other variables as well is clearly needed. Despite these issues in my analysis of the topic however, this study shows that as income inequality increases, mortality also increases at the county level, and with the large sample size and multiple years used in the analysis, this result lends support to previous studies showing the same relationship.

Moving forward, this study has created many research questions which need to be investigated further in future studies. As this analysis is the first ever comprehensive cross-sectional time-series analysis of income inequality and health outcomes at the county level, these results need to be replicated by other researchers before the results of the analysis can be truly trusted. Also, more research is needed to understand if and how presidential vote choice, represented in this analysis by Republican deviation from the national average, actually affects mortality. With the variable there appears to be a paradox between Republicans arguing against government intervention into health despite receiving its benefits, and Republicans having better health outcomes. This outcome is not supported by any previous research, and although potentially explainable

using the argument that Republicans are wealthier and thus have better access to medical care, more research on the subject is clearly needed. Furthermore, the success of the Republican deviation variable in this analysis also brings to light another interesting research question: what other political variables have an impact on health outcomes?

While on the surface it may seem ridiculous that political variables could actually have a direct impact on mortality, with more analysis it becomes clear that politics impacts every aspect of daily life and for that reason it is plausible that several other political variables could impact health. For that reason I recommend the inclusion of new political variables in future research, with government health expenditures, public approval ratings on government intervention into health, and congressional votes on welfare policies as examples of possible variables to use. Lastly, with the land area variable unexpectedly showing mortality to be lower in counties that are larger, more research is needed into the relationship between mortality and geographic attributes. There is little research currently available on the topic, and with both land area and population density proving significant with and without the income control variable, more research on the topic is needed.

Ultimately, this analysis of income inequality and health outcomes has shown that income inequality has a strong positive correlation with mortality at the county level. This study is a new addition to the field as it is the first ever comprehensive cross-sectional time-series analysis done at the county level and is also the first study to my knowledge to include any political variables in their analysis at the sub-national level. Furthermore, this analysis uses a new measure of income inequality which simplifies the formerly arduous process of calculating income inequality and it yields the same results as previous studies of income inequality and health outcomes. Lastly, through its

inclusion and discussion of future political variables for use in analysis, this study helps to advance the social epidemiology literature and connect the seemingly related but rarely associated fields of political science and social epidemiology.

Table 1. Mortality Rates Influences

	(1)
Income Inequality	35.15*** (-10.04)
African-American	0.000331*** (-0.0000647)
Female	-6.03e-05*** (-1.93E-05)
Medicare	1,470*** (-130.4)
Poverty	1.885*** (-0.474)
Population Nat. Log	-50.40*** (2.685)
Unemployment	3.927*** (-0.867)
Mean Age Group	-120.0*** (-10.7)
Republican Deviation	-1.283*** (-0.201)
Population Old	48.55*** (-2.074)
landarea_100miles	-1.352*** (-0.25)
Population Density	-4.893*** (-1.415)
Constant	1,419*** (-69.76)
Observations	14,583
R-sqd	.7491

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 2 Mortality Rate Influence w/ Income

	(1)
Income Inequality	178.7*** (-11.98)
African-American	0.000162*** (-0.0000563)
Female	-0.00000183 (-0.0000244)
Medicare	1,261*** (-119.6)
Poverty	-0.431 (-0.471)
Population Nat. Log	-22.36*** (-3.15)
Unemployment	-0.17 (-0.881)
Mean Age Group	4.236 (-12.96)
Republican Deviation	-0.627*** (-0.19)
Population Old	31.86*** (-2.238)
Income	-0.00384*** (-0.000184)
landarea_100miles	-1.474*** (-0.258)
Population Density	-3.162** (-1.23)
Constant	523.8*** (-83.81)
Observations	14,583
R-sqd	.7718
Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1	

Figure 1: Histogram of Mortality Rates per 100,000 Individuals

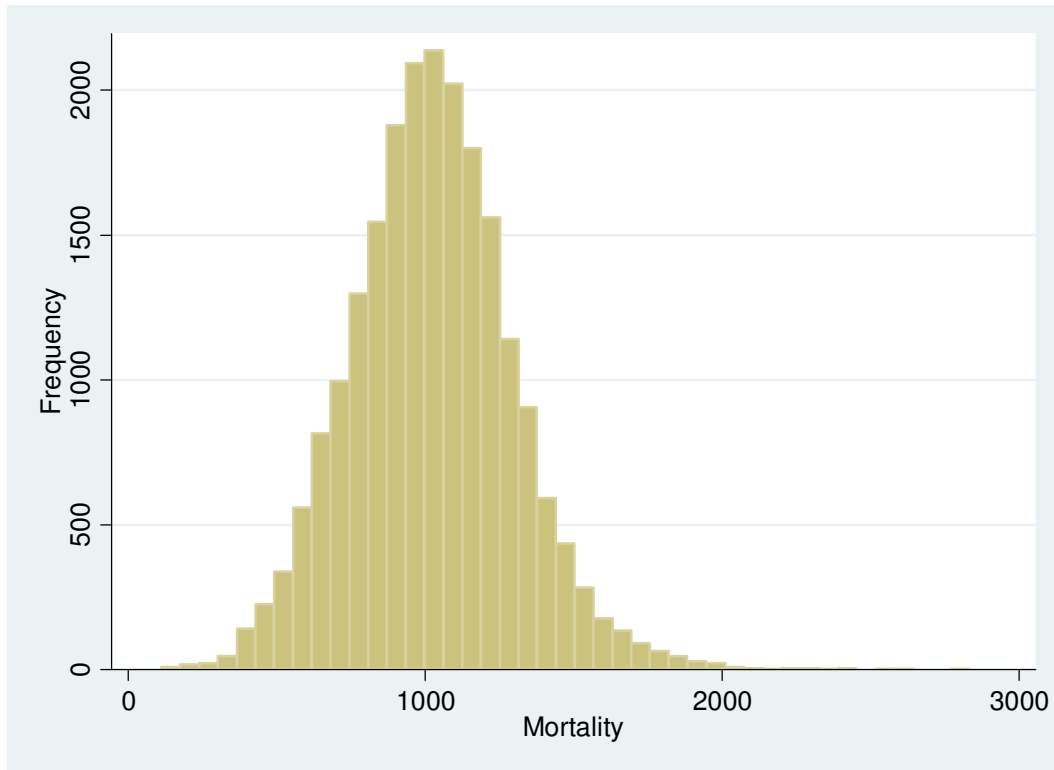


Figure 2: Histogram of Income Inequality Using Mean/Median Income

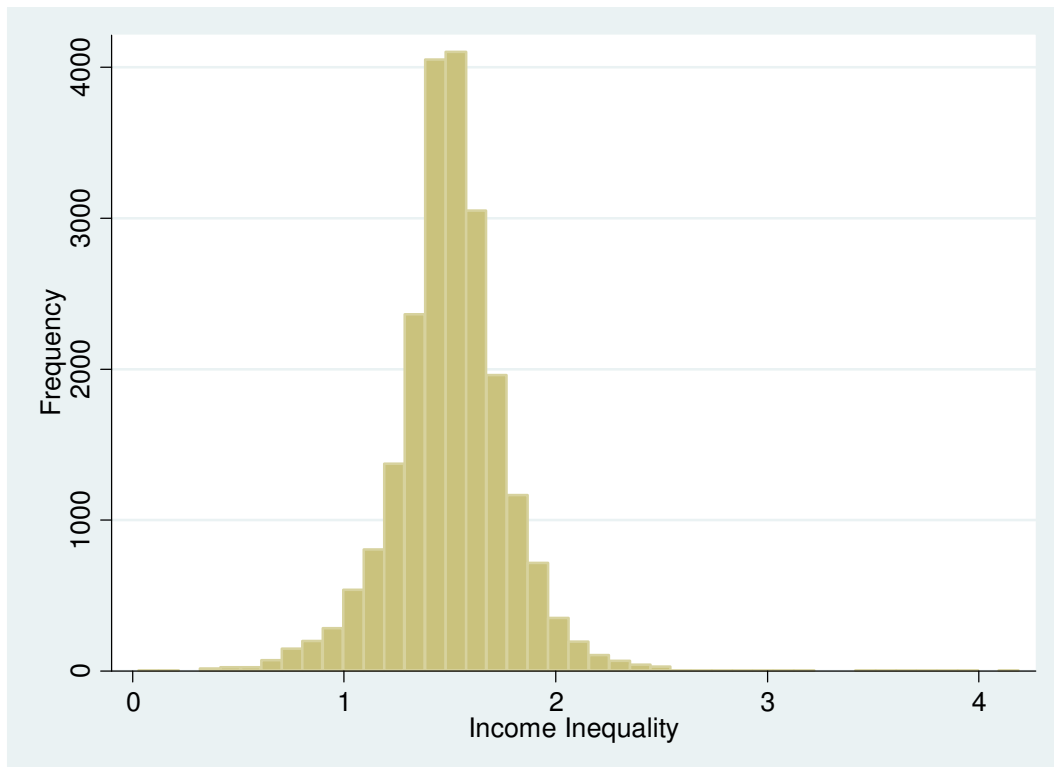


Figure 3: Histogram of Mean Housing Unit Income

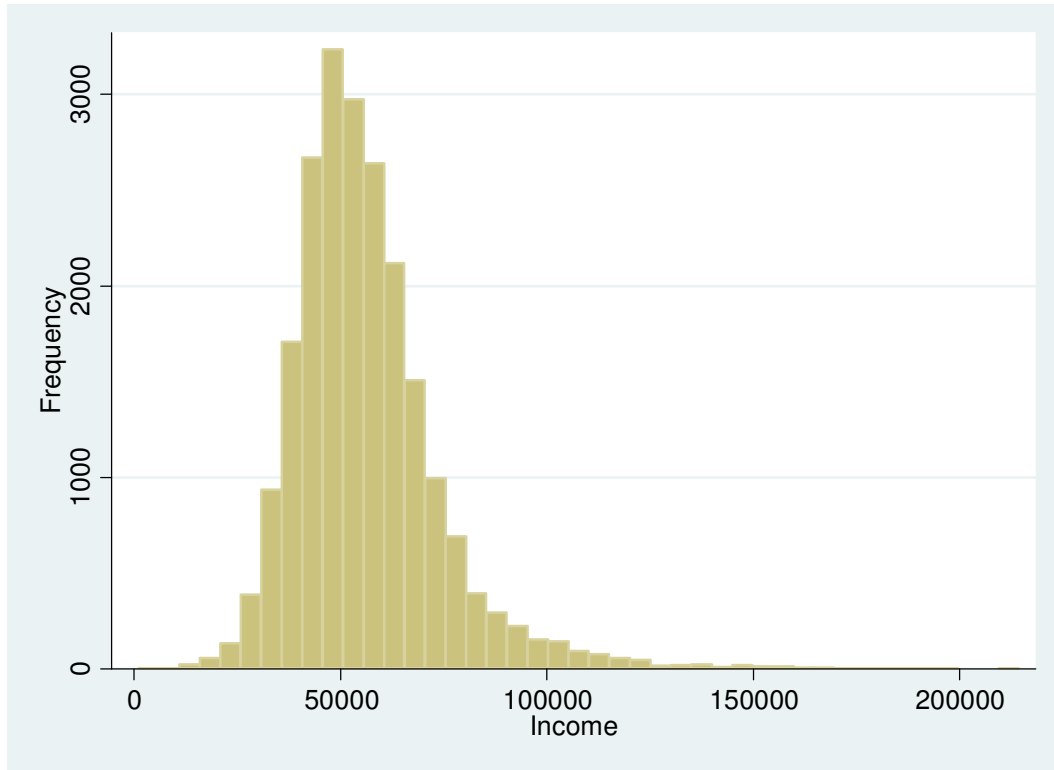


Figure 4: Histogram of the Percentage of the Population under the Poverty Line

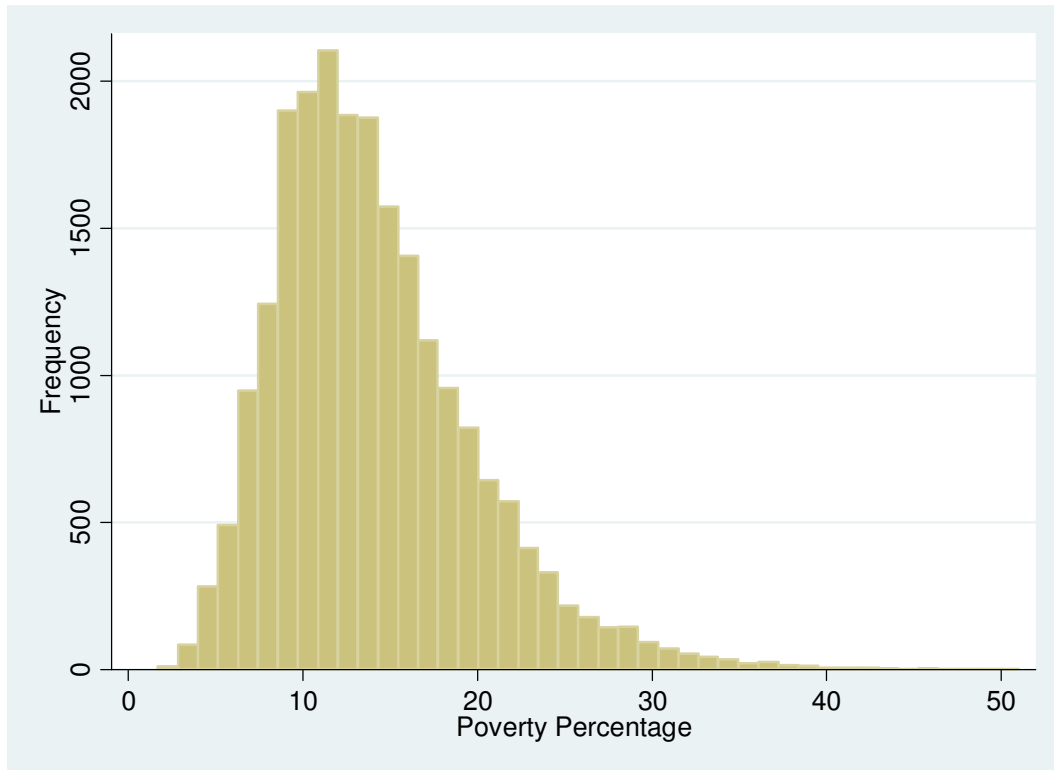
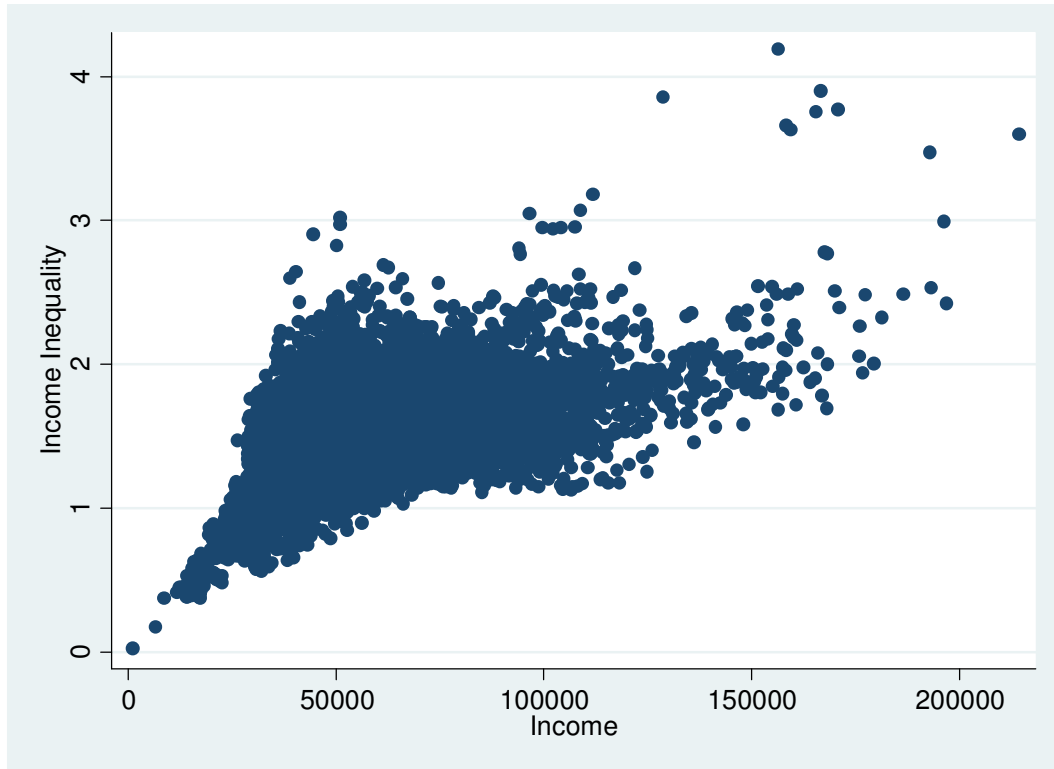


Figure 5: Scatter-plot of Income and Income Inequality



References

- Auger,N., G. Zang, and M. Daniel. 2009. "Community-Level Income Inequality and Mortality in Québec, Canada." *Public Health* 123 (6):438-43.
- Bartels, Larry M.,. 2008. *Unequal democracy : The political economy of the new gilded age*. New York; Princeton: Russell Sage Foundation ; Princeton University Press.
- Beyers, Matt, Janet Brown, and Alameda County Public Health Department (Calif.). 2008. *Life and death from unnatural causes : Health and social inequality in alameda county*, Anonymous Anonymous . Oakland, Calif: CAPE unit of the Alameda County Public Health Department.
- Brodish,P. H., M. Massing, and H. A. Tyroler. 2000. "Income Inequality and all-Cause Mortality in the 100 Counties of North Carolina." *Southern Medical Journal* 93 (4):386-90.
- CDC Wonder Customer Support. Personal Communication, October 6th, 2010.
- Centers for Disease Control and Prevention, National Center for Health Statistics.
Compressed Mortality File 2000-2006. CDC WONDER On-line Database, compiled from Compressed Mortality File 2000-2006 Series 20 No. 2L, 2009. Accessed at <http://wonder.cdc.gov/cmfi10.html> on Nov 1, 2010 10:13 AM
- Centers for Disease Control and Prevention, National Center for Health Statistics.
Compressed Mortality File 1979-1998 and 2000-2006. CDC WONDER On-line Database. <http://wonder.cdc.gov/wonder/help/cmfi10.html>

- Franzini,L., J. Ribble, and W. Spears. 2001. "The Effects of Income Inequality and Income Level on Mortality Vary by Population Size in Texas Counties." *Journal of Health and Social Behavior* 42 (4):373-87.
- Franzini,Luisa, and William Spears. 2003. "Contributions of Social Context to Inequalities in Years of Life Lost to Heart Disease in Texas, USA." *Social Science & Medicine* 57 (10) (Nov):1847-61.
- Gelman, Andrew. 2008. *Red state, blue state, rich state, poor state : Why Americans Vote the Way They Do*. Princeton: Princeton University Press.
- “Getting Started” 2010. Medicare.gov. <http://www.medicare.gov/>
- Kawachi,I., S. V. Subramanian, and N. Almeida-Filho. 2002. "A Glossary for Health Inequalities." *Journal of Epidemiology and Community Health* 56 (9):647-52.
- Kelly, Nathan J. 2009. *The Politics of Income Inequality in the United States*. Cambridge: Cambridge University Press.
- Ladewig, Jeffrey. 2010. “Income Inequality and Ideological Position in the U.S. Congress.” Midwest Political Science Association Conference.
- Ladewig, Jeffrey and Timothy H. Callaghan. 2011. “Death and Politics: The Social, Economic, and Political Sources of County-Level Mortality Rates in the United States.” Midwest Political Science Association Conference.

Lynch, John W., George Davey Smith, George A. Kaplan, and James S. House. 2000.

"Income Inequality and Mortality: Importance to Health of Individual Income, Psychosocial Environment, Or Material Conditions." *BMJ: British Medical Journal* 320 (7243) (Apr. 29):pp. 1200-1204.

Marmot, Michael, and Richard G. Wilkinson. 2001. "Psychosocial and Material Pathways in the Relation between Income and Health: A Response to Lynch Et Al." *BMJ: British Medical Journal* 322 (7296) (May 19):pp. 1233-1236.

McCarty, Nolan M., Keith T. Poole, and Howard Rosenthal . 2006. *Polarized america : The dance of ideology and unequal riches*. Cambridge, Mass.: MIT Press.

Page, Benjamin I., and Lawrence R. Jacobs. 2009. *Class war? : What americans really think about economic inequality*. Chicago; London: University of Chicago Press.

Reitan, Therese C. 2003. "Too Sick to Vote? Public Health and Voter Turnout in Russia during the 1990s." *Communist and Post-Communist Studies*. 36 (1):49.

Soss, Joe, and Lawrence R. Jacobs. 2009. "The Place of Inequality: Non-Participation in the American Polity." *Political Science Quarterly*. 124 (1):95.

"Support Center" 2001. Medicare.gov.

http://questions.medicare.gov/app/answers/detail/a_id/10/session/L3NpZC9kSnpxbnNzaw%3D%3D

- Theodossiou,I., and A. Zangelidis. 2009. "The Social Gradient in Health: The Effect of Absolute Income and Subjective Social Status Assessment on the Individual's Health in Europe." *Economics and Human Biology* 7 (2):229-37.
- Wilkinson,R. G. 1997a. "Socioeconomic Determinants of Health: Health Inequalities: Relative Or Absolute Material Standards?" *British Medical Journal* 314 (7080):591-5.
- Wilkinson,R. G., and K. E. Pickett. 2008. "Income Inequality and Socioeconomic Gradients in Mortality." *American Journal of Public Health* 98 (4):699-704.
- Wilkinson,Richard G. 1997b. "Comment: Income, Inequality, and Social Cohesion." *American Journal of Public Health* 87 (9) (09):1504-6.
- Wilkinson,Richard G., and Kate E. Pickett. 2006. "Income Inequality and Population Health: A Review and Explanation of the Evidence." *Social Science & Medicine* 62 (7) (Apr 2006):1768-84.